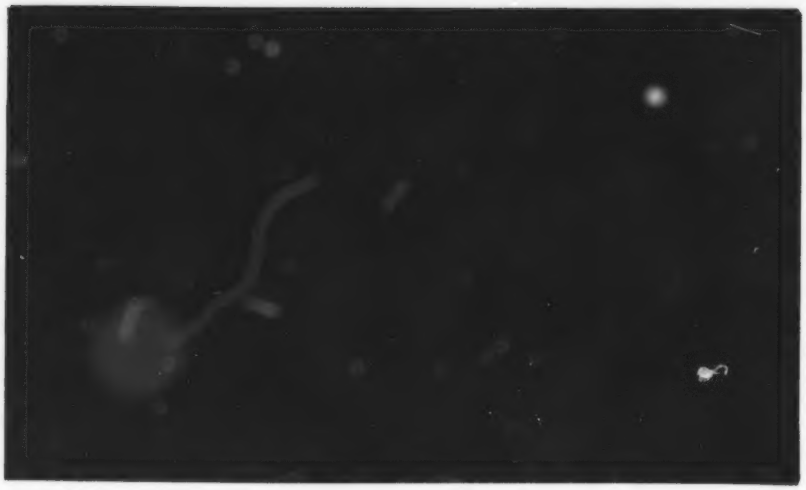
CENTRE
DE RECHERCHE
SUR LES TRANSPORTS
(CRT)
CENTRE
FOR RESEARCH
ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE



C.P. 6128, SUCCURSALE A
MONTRÉAL (QUÉBEC)
CANADA
H3C 3J7

TÉLÉPHONE : (514) 343-7575
TÉLEX : 05-24146 BIBPOLYTEC
TÉLÉCOPIEUR (FAX) : (514) 343-7121

CRT - 649

A DYNAMIC TRAFFIC ASSIGNMENT MODEL AND A SOLUTION ALGORITHM

by

Omar Drissi-Kaïtouni
and
Abdelhamid Hameda-Benchekroun

This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

Université de Montréal
Centre de recherche sur les transports - Publication #649
September 1989

90-01405/

Dépôt légal - 4e trimestre 1989

Abstract

This paper is concerned with the modelling of the Dynamic Traffic Assignment Problem (DTAP) for predicting the flows of urban transportation networks, mainly at peak periods. During the past 40 years, most of the research has been in the Static Traffic Assignment Problem (STAP) where it is assumed that the demand is constant over the time. This assumption is realistic for the analysis of intercity freight transportation networks over long periods of time, but it does not hold in urban area, for simulating the flow variations during short periods (peak hours). Hence, during the past 20 years, the interest to the DTAP has been increasing: in brief, most of the research before 1978 led to heuristic models (where the demand is assigned to *instantaneous* minimal cost paths), and after 1978, global optimization models have been proposed, but all of them incorporate important limitations (only one destination; unrealistic conditions on the cost functions in order that the flow "reaches" the destination; possible violation of the link capacities; etc.).

In this paper, we propose a Dynamic Traffic Assignment Model which is mainly based on the following assumption: the time spent by a vehicle on a link may be decomposed into a fixed travel time plus a waiting time. The fixed travel time corresponds to the free or uncongested travel time over the link. Then the vehicle is put in an exit queue (which resides on the same link) until it becomes possible to enter a forward link; this decision is based on the link costs and their capacities. We show that this model leads to a network structure (a temporal expansion of the base network, including the queues) and therefore the DTAP may be viewed as a "simple" STAP over the expanded network. Hence, all the theory developed during the past 40 years for the STAP may be used to solve the DTAP.

Key words: Dynamic, Traffic, Assignment, Network, Equilibrium.

Résumé

Cet article concerne le problème d'affectation dynamique du trafic pour l'analyse des réseaux de transport urbain aux heures de pointes. En bref, les études effectuées avant 1978 ont abouti à des heuristiques, et après 1978, des modèles d'optimisation ont été proposés, mais ils comportent tous des limitations importantes (une seule destination, conditions non réalistes sur les fonctions de coût, violation possible des capacités des liens, etc.).

Dans cet article, nous proposons un modèle d'affectation dynamique du trafic, principalement basé sur les hypothèses suivantes: le temps passé par un véhicule sur un lien peut être décomposé en un temps de traversée fixe plus un temps d'attente. Le temps de traversée fixe correspond au temps requis pour traverser le lien lorsqu'il n'y a pas de congestion, après quoi le véhicule se met en file d'attente pour sortir du lien. Le temps passé dans la file d'attente dépend des coûts, des volumes et des capacités des liens du réseau. On montre que ce modèle dynamique peut être perçu comme un modèle d'affectation statique sur une expansion temporelle du réseau. Les méthodes classiques utilisées pour résoudre le problème d'affectation statique du trafic s'appliquent alors pour la résolution des problèmes relatifs à notre modèle.

Mots clés: Affectation, Dynamique, Trafic, Réseau, Equilibre.

1- Introduction

This paper deals with the modelling of the Dynamic Traffic Assignment Problem (DTAP) which is aimed at simulating the flows of urban transportation networks, mainly at peak periods. During the past fourty years, most of the research has been concentrated in the Static Traffic Assignment Problem (STAP or simply TAP) where it is assumed that the demand is constant over the time. This assumption is realistic for the analysis of intercity freight transportation networks over long periods of time (as a month or a year) where the traffic flows may be supposed to have reached some steady state. As for a city, the traffic flow over its network of streets is not constant over a day; however the static models have been extensively used in practice in urban contexts. Many static models have been proposed in the literature; see for instance Wardrop [1952] for uncongested networks, Dafermos and Sparrow [1969] for congested networks, and Florian et al. [1987] and Guélat et al. [1987] for multimode - multiproduct congested networks. Under the assumption of uncongestion on the network, the time spent by vehicles on the links is independent of the flow values, hence the DTAP reduces to a sequence of STAP's. But during peak periods, in general, heavy congestion occures on strategic links of the network, and a static model is no more sufficient to explain the traffic flows; this has motivated the study of dynamic models for which the interest has been increasing during the past twenty years.

All the research before 1978, regarding the DTAP, led to heuristic models (see for instance Yagar [1970] or the computer system CONTRAM (Leonard et al.[1978])) where the demand is assigned on instantaneous minimal cost (shortest) paths; this choice is not optimal since it does not take account of the demand over the whole horizon. Other models have been proposed, but all of them incorporate important limitations: Robillard's model [1974] does not apply to capacity constrained links (e.g. congested networks) and Hendrickson and Kocur [1981], Arnott et al. [1987] and De Palma et al. [1987] consider only parallel paths networks where congestion occures at a single link (bottleneck).

In 1978, Merchant and Nemhausser present a first (discret time nonconvex) minimization model for the single origin-destination DTAP. Ho [1980] proposes a stepwise linear version of this model which has been more deeply investigated by Carey [1986 and 1987]; this author presents a convex version of the model and some technics to generalize it to the many origins and many destinations case. Finally, Friesz et al. [1988] propose a continuous time version (optimal control problem).

In the model of Merchant and Nemhausser [1978], and its variants, the authors represent the links' dynamic of flow by:

$$x_a^{t+1} = x_a^t + d_a^t - g_a(x_a^t) \quad \forall a \in A, t = 0, 1, .., T-1$$

where:

x_a^t and x_a^{t+1} are the flows on link a at periods t and $t+1$,
d_a^t is the flow that enters link a at period t,
$g_a(x_a^t)$ is the flow that exits link a at period t,
A is the set of links of the network, and
T is the number of periods of study (horizon).

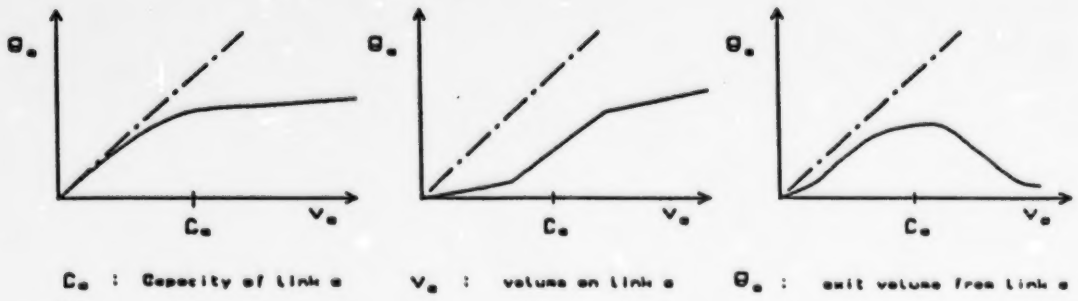


Figure 1: Different types of exit functions

The flow that exits link a is given by mean of an "exit function" $g_a(.)$ which depends only on the link flow; many types of functions (see figure 1) have been proposed for the $g_a(.)$'s by the authors of this model and its variants, but none of them seems to be entirely satisfactory for the following reasons:

- first, all the proposed exit functions $g_a(.)$ (for a link a) are independent of the flows and capacities of the others links of the network, which is not realistic since the flow that exits a link may depend in major part on the forward links' flows and capacities (for instance, the links preceeding a bottleneck are rather influenced by the bottleneck than by their own flow),
- second, the flow $g_a(x_a^t)$ that exits a link a may violate the capacity of the links in which it enters during the next period,
- third, the use of continuous exit functions presumes that the flow on each link, at each period, is uniformly distributed on the link; this assumption does not necessary hold in a dynamic system, mainly during the periods of uncongestion and the transition periods between uncongestion and congestion.

Apart the difficulties mentioned above, the dynamic flow conservation constraints of the model do not ensure that the flow will reach the destination (independently of the horizon T); in order to come over this problem, the authors of the model impose an unrealistic condition on the link cost functions, that is roughly that the cost of the links must increase with their distance to the destination.

A final lacuna, which is not the least, is that the induced mathematical programs have not a special structure that may lead to a particularly efficient adaptation of mathematical programming algorithms for the solution of the problem.

Long time ago, an important model was proposed for the Dynamic Maximal Flow Problem by Ford and Fulkerson [1962] and Fulkerson [1975]. In this model, each link $a \in A$ has a capacity and a fixed traversal time. For any period of time, at each node vehicles can either be trans-shipped on exiting links or wait until the next period. The resulting space-time

network is built as follows:

- each node $n \in N$ is expanded to $T+1$ nodes n_t, $t = 0, 1, .., T$, and
- each link $a = (i, j) \in A$ is expanded to the links $(i_t, i_{t+\tau_a})$ where τ_a is the fixed traversal time on link a and t such that $t \geq 0$ and $t + \tau_a \leq T$; we refer to these links as travel links.
- the waiting time at the nodes is represented by mean of the links (n_t, n_{t+1}) where $n \in N$ and $t = 0, 1, .., T-1$; we refer to these links as the queues.

Figure 3 illustrates an expansion of the network of figure 2, according to the model described above.

Figure 2: A simple network

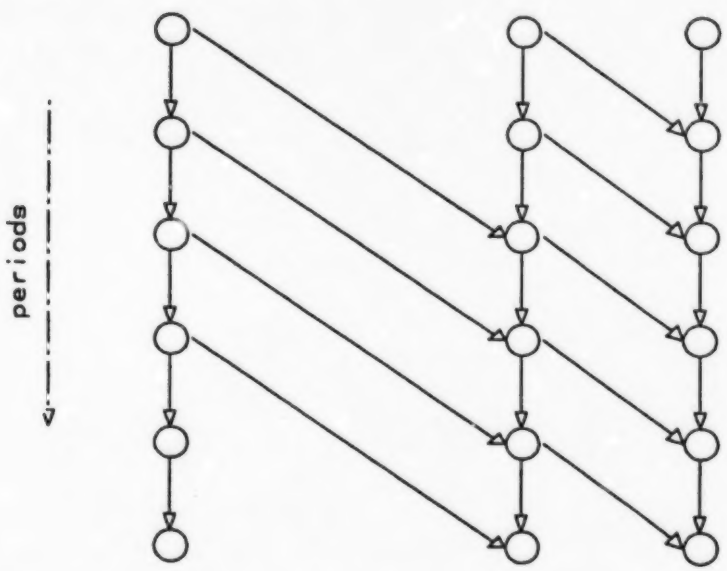


Figure 3: Expansion of the simple network

This model predicts that when there is no congestion on the network, the time spent on each link is constant (free or uncongested traversal time), and, in presence of congestion, it increases because of waiting in queues; this meets our expectation. The major objection to the model is that queues hold at the nodes rather than on the links, and no capacity is associated with the queues volumes. This means that the nodes of the city network (intersections) are viewed as infinite stacks for the vehicles.

Recently, Zawack and Thompson [1987] adapted this model for the DTAP with the terminology of green links for the travel links and red links for the queues. However, these authors brought some improvements to the space-time expanded network of Ford and Fulkerson [1962] and Fulkerson [1975]:

- finite capacities are set on the queues, and
- traversal links and queues are not allowed in the expanded network if there is no space-time path connecting them to an origin node; this helps to reduce the size of the expanded network.

These authors formulate the DTAP as a trans-shippement problem over their expanded network where they consider only one destination and many origins.

However, although these improvements have been brought, setting the queues at the nodes remains a major problem because it is not clear, and in fact not even possible, to characterize adequately the node queues capacities. In fact, the queues are physically on the links (streets) rather that on the nodes (intersections), and the queue on a link shares the capacity of the link with the other vehicles on the same link (but not yet in the queue). Setting the queues at the nodes, as in the models above, means that the queues of all the links sharing a same head node are grouped together to form a single node queue. But each of the components (link queues) of a node queue has its own capacity that it does not share with the other components. As a consequence, it is not possible to define global fixed node queues capacities, as suggested by Zawack and Thompson [1987]. Our point of view is that a more realistic model must incorporate link queues rather than node queues.

The remainder of this paper is organized as follows: in the next section, we describe a space-time dynamic network in the spirit of those of Ford and Fulkerson [1962], Fulkerson [1975] and Zawack and Thompson [1987], where we incorporate explicit link queues, and then we state our Multi-Commodity (many origins and many destinations) Dynamic Traffic Assignment Model. In section 3, we propose a solution algorithm, and we conclude the paper in section 4.

2- The model

We begin this section by stating the assumptions underlying our model and by describing our Dynamic Flow Conservation Constraints (DFCC). Hence, we describe the corresponding temporal expanded network. Finally, we present the whole mathematical model.

2.1- Assumptions

The main assumptions underlying our model are the following:

-A1- For any period of time the (nonnegative) link flows are bounded (from above) by the link capacities; hence, a vehicle cannot enter a congested link (e.g. at full capacity); it

must wait (on the link in which it is) until that a place becomes free in the congested link.

-A2- The total time spent by a vehicle on a link a is equal to a :
- fixed travel time τ_a which is the free (uncongested) travel time on link a, after which the vehicle is ready to exit the link (e.g. the vehicle joins the exiting queue of the link), plus a
- waiting time which is the time spent in the exiting queue of the link; according to A1, a vehicle has to wait if the link in which it must enter is congested (e.g. at full capacity).

-A3- The free travel time τ_a of link a is measured in numbers of periods of time.

-A4- At the initial time period t=0, the flow is null on all the links, and so it is at the last time period $t = T$ (e.g. all the vehicles have reached their destination at t= T).

This last assumption, which is maid here for simplifying the presentation, is not really restrictive and has always been implicitly supposed in the literature. However, we shall see later how to relax it in order to incorporate into the model the presence of flow at the initial period, and also at the last period (vehicles that have not yet reached their destination, because of a too short study horizon).

In the next subsection, we describe the dynamic of the transfer of the flow between and in the links in our model.

2.2- Dynamic Flow Conservation Constraints (DFCC)

Let A be the set of links of the network, and T the horizon (number of periods). Thus, let's define, for each $a \in A$ and each time period t:
x_a^t : total flow on link a at the begining of period t,
u_a^t : the flow that enters in link a during period t,
v_a^t : the flow that exits from link a during period t,
q_a^t : part of x_a^t, ready to exit from link a at the begining of period t [1]

The travel time being τ_a for link a, the flow $u_a^{t-\tau_a}$ is ready to leave the link a at the begining of period t; in other words, $u_a^{t-\tau_a}$ joins the queue q_a^t during period t, and since v_a^t leaves the link during period t, the DFCC for link a at period t is:

$$q_a^{t+1} = q_a^t + u_a^{t-\tau_a} - v_a^t \quad \forall a \in A. \tag{1}$$

In a similar way, let F_k^t the demand at origin node k, $k \in O$, during period t. F_k^t joins the queue Q_k^t at node k during period t, and let b_k^t the part that enters the network during this period, then the DFCC for the input flow is:

$$Q_k^{t+1} = Q_k^t + F_k^t - b_k^t \quad \forall k \in O. \tag{2}$$

[1] We note that all the queues are empty at $t = 0$ (assumption A4), and $u_a^{t-\tau_a} = 0$ for $t - \tau_a \leq 0$.

The transfer of flow at the nodes must satisfy that the arriving flow is equal to the exiting flow at any period of time t, which is expressed by:

$$\sum_{a \in B_k} v_a^t + b_k^t = \sum_{a \in A_k} u_a^t + e_k^t \quad \forall k \in N \tag{3}$$

where

B_k is the set of links arriving to the node k,
A_k is the set of links exiting from the node k,
$b_k^t = 0$ if $k \notin O$,
$e_k^t = 0$ if $k \notin D$ where:
D is the set of destination nodes, and
e_k^t is the flow that exits at destination node k during period t.

Equations (1), (2) and (3) represent the DFCC of our model. Now we turn to the corresponding static network.

2.3- The Static Temporal Expanded Network (STEN)

In this subsection, we construct the temporal expanded network $\mathcal{G} = (\mathcal{A}, \mathcal{N})$ that corresponds to the flow conservation constraints of the preceeding subsection.

Let θ_k^i be the length of the shortest path from the origin node n_i to the node n_k, and let $\theta_k^* = min_{i \in O} \theta_k^i$, which corresponds to the minimal time required to reach the node n_k from any origin; for $t < \theta_k^*$ no flow enters node n_k.

• Nodes: each node n_k, $(k \in N)$ of the base network is expanded to $(T+1)$ nodes n_k^t, $t = 0, 1, .., T$, $(k \in N)$ of the expanded network.

• Links: for each link $a = (n_i, n_j)$, $\in A$, we construct, for each $t = 0, 1, .., T$, such that $t - \tau_a \geq \theta_i^*$:
- one node l_a^t,
- one link $(n_i^{t-\tau_a}, l_a^t)$,
- one link (l_a^t, n_j^t) and
- one link (l_a^t, l_a^{t+1}) if $t < T$.

These three links correspond to the variables $u_a^{t-\tau_a}$, v_a^t and q_a^{t+1}, and constraints (1) correspond to the (static) flow conservation constraints at the nodes l_a^t, $a \in A$, $t = 0, 1, .., T$, such that $t - \tau_a \geq \theta_i^*$:

• Origin nodes: for each origin node n_k, $k \in O$, we construct, for each $t = 0, 1, .., T$:
- one node $L_{n_k^t}$,
- one link $(L_{n_k^t}, n_k^t)$ and
- one link $(L_{n_k^t}, L_{n_k^{t+1}})$ if $t < T$.

These two links correspond to the variables b_k^t and Q_k^{t+1}, and constraints (2) correspond to the (static) flow conservation constraints at the nodes $L_{n_k^t}$, $k \in O$, $t = 0, 1, .., T$.

• Destination nodes: for each destination node n_k, $k \in D$, we construct:
- one node d_k and
- one link (n_k^t, d_k) for each $t = 0, 1, .., T$, $t \geq \theta_k^*$.

These links correspond to the variables e_k^t where d_k is a super-destination for all the periods for the destination node n_k.

In order to illustrate this, we consider the simple network of figure 2, and show the corresponding static expanded network (for five periods) in figure 4.

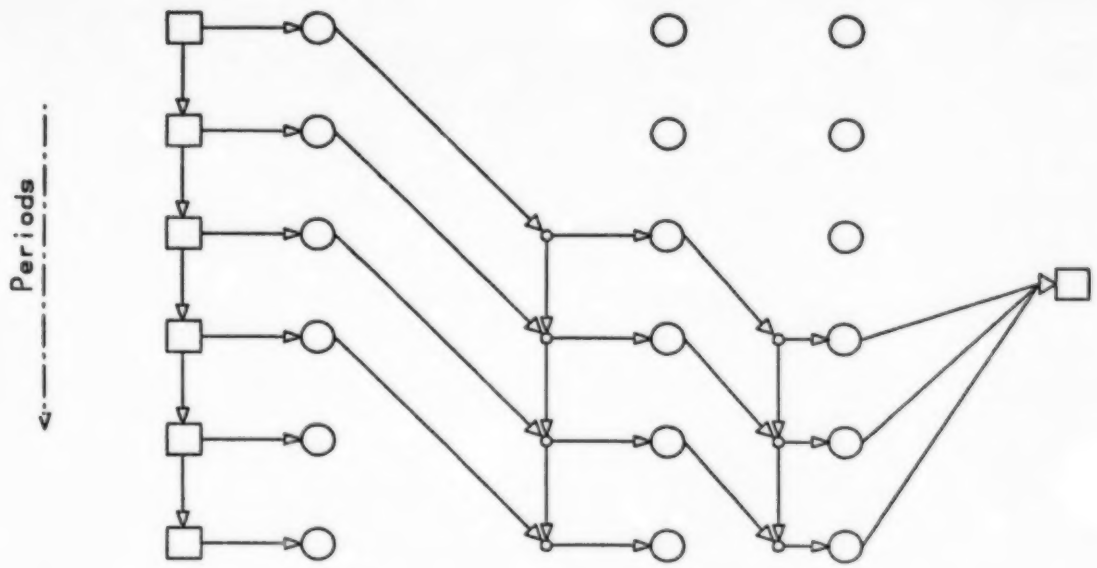


Figure 4: Expansion of the simple network of Fig. 2

The maximal size of the STEN, for T periods of study, is given in table 1:

	Base network	STEN
Number of origins	O	O*T
Number of destinations	D	D
Number of links	A	(3*A + 2*O + D)*T
Number of nodes	N	(N + O + A)*T + D

Table 1

Now we turn to the flow bounds on the links. Since the nonnegativity constraints apply of course to all the variables, we turn immediately to the capacity constraints without which all the queues would always be empty and the dynamic problem would reduce to a sequence of static ones.

2.4- Capacity constraints

• Upper bound on the total link flow: at each period t, the total flow x_a^t on the link a, must not exceed $\tau_a C_a$, where C_a is the capacity of link a, expressed in vehicles/period. Hence,

$$x_a^t \leq \tau_a C_a \quad \forall a \in A, \quad t = 0, 1, .., T.$$

x_a^t is all the flow that entered but has not yet left link a at period t; in other words, it is equal to the present queue q_a^t plus the flow on link a that has not yet joined the queue, which is mathematically expressed by: $x_a^t = q_a^t + \sum_{i=t-\tau_a}^{t-1} u_a^i$.

Hence, the total link flow capacity constraints take the form:

$$q_a^t + \sum_{i=t-\tau_a}^{t-1} u_a^i \leq \tau_a C_a \quad \forall a \in A, \quad t = 0, 1, .., T. \tag{4}$$

• Upper bound on the entering flow: the maximum flow that can enter a link during one period is the link capacity C_a (expressed as a number of vehicles/period):

$$u_a^t \leq C_a \quad \forall a \in A, \quad t = 0, 1, .., T. \tag{5}$$

• Upper bound on the exiting flow: we have the similar constraints for the flow that exits the links:

$$v_a^t \leq C_a \quad \forall a \in A, \quad t = 0, 1, .., T. \tag{6}$$

The set of constraints (4), (5) and (6) represent the capacity constraints of our model.

2.5- The mathematical model

In this subsection, we consider the STEN $\mathcal{G} = (\mathcal{A}, \mathcal{N})$ described in subsection 2.3. Let $g_{n_p^t d_q}$ be the demand at period t at the origin node n_p for the destination node n_q (d_q in the expanded network); $g_{n_p^t d_q}$ must be assigned on the paths $k \in K_{pq}^t$, where K_{pq}^t is the set of paths between the nodes n_p^t and d_q in the STEN $\mathcal{G}$. We consider a user equilibrium assignment since each user seeks to minimize its own travel cost. Hence, the problem is a "simple" static user equilibrium problem over the STEN $\mathcal{G}$, where, however, the link flows must satisfy the capacity constraints (4) to (6). Therefore, the model that we propose takes the following form:

$$min \sum_{a \in \mathcal{A}} \int_0^{f_a} s_a(z)dz$$

$$s.t. \sum_{k \in K_{pq}^t} h_k = g_{n_p^t d_q} \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T,$$

$$h_k \geq 0 \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T, \ \forall k \in K_{pq}^t$$

$$f_a = \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} \delta_{ak} h_k \ \ \forall a \in \mathcal{A}$$

$$\delta_{ak} = \begin{cases} 1 & \text{if link } a \text{ is in path } k \\ 0 & \text{otherwise} \end{cases}$$

and the capacity constraints (4), (5) and (6)

where, h_k is the flow on path $k \in K$, f_a the flow on link $a \in \mathcal{A}$. Now, we would like to concentrate on some important details of the model:

2.5.1- Cost functions

The transportation unit cost on a link usually refers to the transportation time. However, it may also include other cost characteristics such as fuel consomption. When there is no congestion on the link, in general, this unit cost is a constant related to the length and physical characteristics of the link. But when congestion is present, the unit cost includes also a delay unit cost which depends on the volumes and capacities of the links of the network; in our model, this delay unit cost may naturally be associated with the time spent in queues. Hence, we suggest to use the following constant unit cost functions:

-travel time on links	(for the u_a^t 's)	:	$\alpha_a \tau_a$
-waiting time on queues	(for the q_a^t 's)	:	β_a
	(for the Q_k^t 's)	:	γ_k

where α_a, β_a, γ_k are time-money conversion factors that may all be equal, but we suggest to use different factors for travel and waiting times; therefore $\alpha_a = \alpha$, $\beta_a = \beta \quad \forall a \in A$, $\gamma_k = \beta \quad \forall k \in N$ and β may be higher than α.

The corresponding (user optimal) objective function would then be:

$$min \sum_{a \in A} \sum_{t=0}^{T} \int_0^{u_a^t} \alpha \tau_a dz + \sum_{a \in A} \sum_{t=0}^{T} \int_0^{q_a^t} \beta dz + \sum_{k \in O} \sum_{t=0}^{T} \int_0^{Q_k^t} \beta dz \qquad (7)$$

but, since the unit costs are constants, the model is also system optimal because the objective function may also be rewritten:

$$min \sum_{a \in A} \sum_{t=0}^{T} \alpha \tau_a u_a^t + \sum_{a \in A} \sum_{t=0}^{T} \beta q_a^t + \sum_{k \in O} \sum_{t=0}^{T} \beta Q_k^t \qquad (8)$$

2.5.2- The fixed tavel times

The fixed travel time τ_a on link $a = (i, j)$ is the mean time (measured in periods) necessary for a vehicle to cross the link from the tail node i to the head node j, when there is no congestion on the link. This time comprises the "normal" delays caused by the stops and red lights on the link, since these are "physical" link characteristics. In order not to count these delays twice, by convention, the delay caused by the stop or red light (if any) of the tail node of the link is not counted in the fixed travel time of the corresponding links.

As a consequence, the fixed travel time is not simply the time to cross the link, corresponding to a constant speed; it should be measured in situ (during a period of noncongestion).

2.5.3- The length of time periods

The base period of time is the unit in which is measured the τ_a 's. Therefore, it must be small enough to ensure that all the τ_a 's are at least equal to 1. The drawback is that if this time period is too small, the number of periods necessary to study the peak hours may be too large. A good compromise is to fix it to a few minutes and to group the shortest links together in order to form links of reasonable lengths.

2.5.4- The O/D demand matrices

In the model, it is assumed that at each period t (which may of length as small as a few minutes), a demand matrix g^t is known. In the reality, it is not reasonable to assume that such matrices may be collected for each of the periods of the study. In general, one matrix for the whole peak hours, or at most one matrix for each of the peak hours may be available. Hence this (these) matrice(s) must be distributed on the periods of the study. Figure 5 shows two ways of doing so.

2.5.5- Relaxation of assumption A4

Initial nonnull flow: If the network has only one destination n, then the initial flow, say x_a^{-1} on link $a = (n_i, n_j)$ may be added to the initial demand for the single destination n, $g_{n_j^0 n}$ at node j. Hence, $g_{n_j^0 n}$ is replaced by $g_{n_j^0 n} + \sum_{a \in A_j} x_a^{-1}$. When there are many destinations on the network, it is not known a priori to which destinations is directed the flow on the links. A simple way to turn over the problem, is to split x_a^{-1} between the $g_{n_j^0 n_q}$'s, $q \in D$. Hence, the $g_{n_j^0 n_q}$'s, $q \in D$ are replaced by:

$$g_{n_j^0 n_q} + \sum_{a \in A_j} x_a^{-1} \frac{g_{n_j^0 n_q}}{\sum_{q' \in D} g_{n_j^0 n_{q'}}}.$$

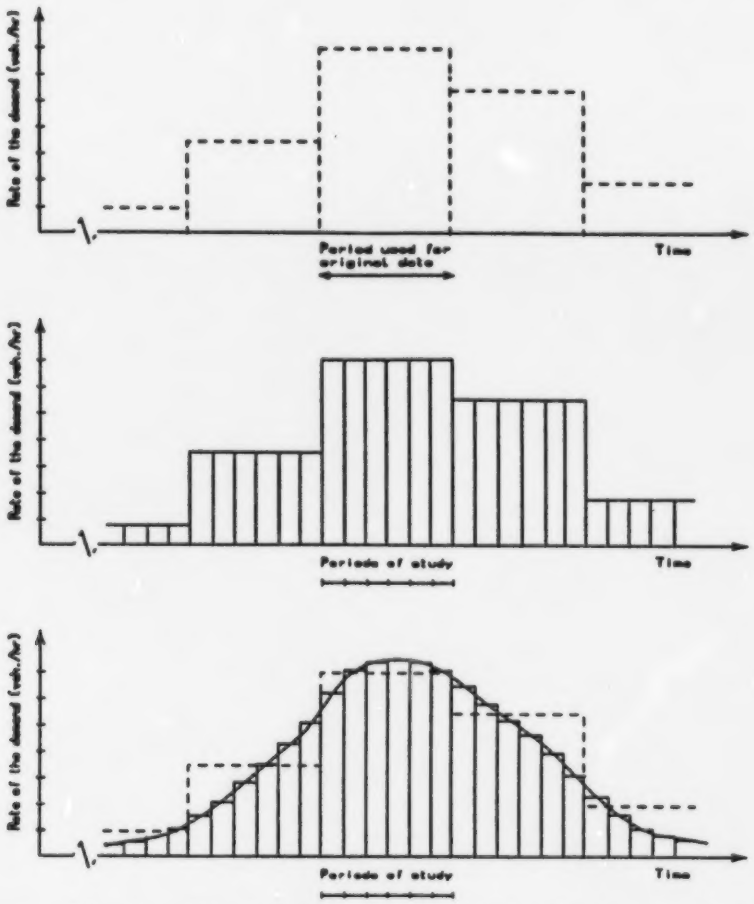


Figure 5: Examples of distribution of the demand on the study periods

Nonnull final flow: The model that we propose is mainly aimed at the analysis of the peak periods. Hence, it is reasonable to assume that the last period T corresponds to the end of peak demand, e.g. of congestion. Therefore, we consider here the case where at the last period T there is still some flow on the network, but not important enough to create congestion. As a consequence, for the periods $t \geq T$, the costs of the links (and paths) do not change in time. Hence, in the STEN, each link $a = (n_i^t, n_j^{t+\tau_a})$ such that $t + \tau_a \geq T$ is replaced by the links (n_i^t, d_q), $q \in D$. The corresponding travel time should be $C_{iq}^t = C_{iq}$ where C_{iq} is the shortest travel time from node n_i to the destination n_q . Hence, a vehicle assigned by the model to a link (n_i^t, d_q) would be on the corresponding physical link (n_i, n_j) at the beging of period $T + 1$. Figure 6 illustrates this for the example network of figure 4.

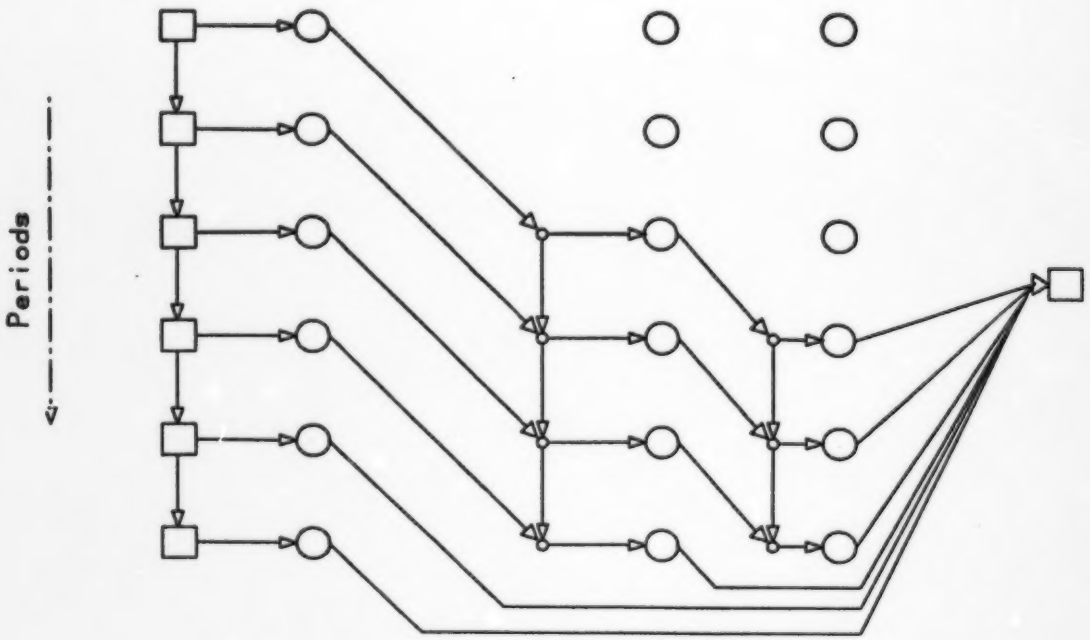


Figure 6: STEN with the links for nonnull final flows

3- Solution of the problem

Usually, in the STAP's, the link flow capacity constraints are incorporated in the objective function as nonlinear unit delay cost functions. This has the effect to simplify the computations when the Linear Approximation method (Frank and Wolfe [1956]) is used to solve the problem, since this algorithm reduces to a sequence of minimal cost paths computations and 1-directional minimizations (see Guélat et al. [1987] for instance).

The approach that we propose to use for the solution of our model is to incorporate the capacity constraints (4) to (6) in the objective in the form of a penality term $PT(f)$, and to solve the resulting problem by a Linear Approximation method; the penality term corresponding to the capacity constraints (4)-(6), may be, for instance,

$$\sum_{a \in A} \sum_{t=0}^{T} \mu[\max(0, u_a^t - C_a)^2 \; + \; \max(0, v_a^t - C_a)^2 \; + \; \max(0, x_a^t - \tau_a C_a)^2]$$

where μ is a large constant.

Therefore, the problem that we consider in the following is:

$$min \sum_{a \in \mathcal{A}} \int_0^{f_a} s_a(z)dz \quad + \quad PT(f)$$

$$s.t. \sum_{k \in K^t_{pq}} h_k = g_{n^t_p d_q} \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T,$$

$$h_k \geq 0 \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T, \ \forall k \in K^t_{pq}$$

with the implicit constraint that :

$$f_a = \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K^t_{pq}} \delta_{ak} h_k \ \forall a \in \mathcal{A}$$

where $$\delta_{ak} = \begin{cases} 1 & \text{if link } a \text{ is in path } k \\ 0 & \text{otherwise.} \end{cases}$$

In a compact form, we seek to solve the problem (P):

$$(P): \ minF(f) \ \ s.t. \ f \in \Omega$$

where Ω is a polytope (linear constraints), which is nonempty[2], and compact as well if the cost functions on all the links of $\mathcal{A}$ are strictly positive (Guélat et al. [1987])[3]. Hence we suggest to adapt to (P) the most widely used algorithm for large scale STAP's, which is the Linear Approximation algorithm (Frank and Wolfe [1956]) that may be stated as:

Linear Approximation Algorithm

Step 0 Initialization
Determine $f \in \Omega$

Step 1 Major Iteration
1.1- Determine extremal point w
1.2- Compute descent direction $d = w - f$
1.3- Compute optimal step size λ
1.4- Update flows $f = f + \lambda d$

Step 2 Stopping Criterion
If $\|\lambda d\| < \epsilon$ then stop
otherwise return to step 1.

[2]The vehicles reach their destination, hence there exist at least one path from each origin to each destination in $\mathcal{G}$, on which all the demand may be assigned; recall that the capacity constraints have been moved into the objective function as a penality term.

[3]The paths considered do not contain cycles.

In step 1.1 of the algorithm, the extremal point w is the optimal solution of the linear approximation of (P) at the current solution f; it solves:

$$(LP): \quad min F(f) + \nabla F(f)^T (w - f) \quad s.t. \quad w \in \Omega$$

where $\nabla F(f)$ is the gradient of F evaluated at the current solution f. This problem is equivalent to:

$$min \nabla F(f)^T w \quad s.t. \quad w \in \Omega$$

or

$$min \sum_{a \in \mathcal{A}} \frac{\partial F(f)}{\partial f_a} w_a \quad s.t. \quad w \in \Omega$$

but $w_a = \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} \delta_{ak} y_k \quad \forall a \in \mathcal{A}$, where y is the vector of path flows corresponding to w. Hence, the objective function of (LP) may be rewritten:

$$min \sum_{a \in \mathcal{A}} \frac{\partial F(f)}{\partial f_a} \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} \delta_{ak} y_k$$

or

$$min \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} y_k \sum_{a \in \mathcal{A}} \frac{\partial F(f)}{\partial f_a} \delta_{ak}$$

or

$$min \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} y_k F_k'(f)$$

where $F_k'(f)$ denotes the marginal cost of path k. Hence, since y is constrained by:

$$\sum_{k \in K_{pq}^t} y_k = g_{n_p^t d_q} \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T,$$

$$\text{and} \qquad y_k \geq 0 \quad \forall p \in O, \ \forall q \in D, \ t = 0, 1, .., T, \ \forall k \in K_{pq}^t$$

the optimal solution is clearly obtained by an all or nothing assignment of the demand on the minimal cost paths (minimal according to the marginal costs $\frac{\partial F(f)}{\partial f_a}$ on the links); the corresponding extremal point in terms of link flows w is obtained by:

$$w_a = \sum_{p \in O} \sum_{q \in D} \sum_{t=0}^{T} \sum_{k \in K_{pq}^t} \delta_{ak} y_k \quad \forall a \in \mathcal{A}$$

where $$\delta_{ak} = \begin{cases} 1 & \text{if link } a \text{ is in path } k \\ 0 & \text{otherwise.} \end{cases}$$

4- Conclusion

This paper is concerned with the modelling of the Dynamic Traffic Assignment Problem (DTAP) for predicting the flows of urban transportation networks during the peak periods.

During the past twenty years, many models have been proposed in the literature, but all of them incorporate important limitations; many of them are heuristic models where the demand is assigned on instantaneous minimal cost (shortest) paths (Yagar [1970], Leonard et al. [1978]). Some others consider only the case of parallel paths between one origin-destination pair (Arnott et al. [1987]). A more realistic model, proposed by Merchant and Nemhauser [1978], which has been deeply investigated by Ho [1980] and Carey [1986], presumes that, for each link of the network, there exist an exit function that determines the number of vehicles that exit of the link as a function of only the link flow. Apart the fact that these functions are to be calibrated in situ for each link, it is not realistic to assume that the exiting flow of a link depends only on the flow of the corresponding link since it may also depend on the capacity and flow of other links of the network. On the other hand, the use of such functions does not ensure that the links capacities will always be satisfied, since the flow exiting a link does not depend on the capacity of the link it enters.

Long time ago, an important model was proposed by Ford and Fulkerson [1962] and Fulkerson [1975] for the dynamic maximal flow problem that has been adapted by Zawack and Thompson [1987] to the DTAP; this model is based on the assumption that the time spent by a vehicle on a link may be separated in two parts:

- a fixed traversal time, and
- a waiting time at the head node.

The fixed traversal times refer to the travel times when no congestion is present, and the waiting times are related to the delays induced by finite capacities on the links during congestion periods. This model leads to a static assignment model over a temporal expansion of the base network that may include many origins, but only one destination.

The main objection to this model concerns the fact that the queues (waiting vehicles) are set on the nodes, which means that a node (e.g. intersection) is considered to be a common stack for the waiting vehicles exiting from all the incoming links to that node.

Our point of view is that a more realistic model should reflect that the waiting vehicles are physically on the links, and, eventually, create congestion on these links (hence, preventing new arrivals into these links), rather than that they wait on the head nodes and eventually permitting new arrivals to the physically congested links.

The model that we propose in this paper is also based on the assumption that the time spent by vehicles on the links may be decomposed into a fixed travel time plus a waiting time in an exiting queue, but the queues are set on the links, and the vehicles in the queue of a link share the physical space with the other vehicles on the link.

Hence, we express the model as a Static Traffic Assignment Promlem (over a temporal expansion of the base network) in terms of path flows allowing the possibility to include many origins and many destinations in the model.

References

ARNOTT, R., de PALMA, A. and LINDSEY, R. (1987) "Departure Time and Route Choice for Routes in Parallel", Discussion Paper #658. Institute for Economic Research.

CAREY, M. (1986) "A Constraint Qualification for a Dynamic Traffic Assignment Model", Operations Research 35 No.5, 55-58.

CAREY, M. (1987) "Optimal Time-Varing Flows on Congested Networks", Operations Research 35 No.5, 58-69.

DAFERMOS, S. and SPARROW, F.T. (1969), "The Traffic Assignment Problem for a General Network", Journal of Research of the National Bureau of Standards 73B, 91-118.

De PALMA, A., HANSEN, P. and LABBÉ, M. (1987), "Commuters' Paths with Penalities for Early or Late Arrival Time", Discussion Paper #8712. Centre for Operations Research & Econometrics, Université Catholique de Louvain.

FLORIAN, M. CRAINIC, T.G. GUÉLAT, J. and Babin, A. (1987), "STAN : An Interactive-Graphic System for Strategic Transportation Analysis at the National and Regional Levels", publ. #495, Centre de Recherche sur les Transports, Université de Montréal.

FORD, L.R. and FULKERSON, D.R. (1962), "Flows in Networks", Princeton University Press. Princeton, New Jersey.

FRANK, M. and WOLFE, P. (1965), "An Algorithm for Quadratic Programming", Naval Reseaech Logistics Quarterly 3:95-110.

FRIESZ, T.L., LUQUE, J., TOBIN, R.L. and WISE, B. (1988) "Dynamic Network Traffic Assignment considered as a Continuous Time Optimal Control Problem", Department of Operations Research and Regional Science, University of Pennsylvania, Philadelphia, PA 19104.

FULKERSON, D.R. (1975), "Stadies in Graph Theory,part I", Studies in Mathematics 11, The Mathematical Association of America.

GUÉLAT, J. FLORIAN, M. and CRAINIC, T.G. (1987), "A Multimode Multiproduct Network Assignment Model for Strategic Planning of Freight Flow", publ. #549, Centre de Recherche sur les Transports, Université de Montréal.

HENDRICKSON, C. and KOCUR, G. (1981) "Schedule Delay and Departure Time Decision in a Determinisic Model", Transportation Science 15, 62-77

HO, J.K. (1980), "A Successive Linear Optimization Approach to the Dynamic Traffic Assignment Problem", Transportation Science 14, 295-305.

LEONARD, D.R., TOUGH, J.B. and BAGULEY, P.C. (1978) "CONTRAM : A Traffic Assignment Model for Predecting Flows and Queues during Peak Periods", Transport and Road Research Laboratory, Report 841, Crowthorne.

MERCHANT, D.K. and NEMHAUSER, G.L. (1978), "A Model and an Algorithm for the Dynamic Traffic Assignment Problem", Transportation Science 12, 183-199.

MERCHANT, D.K. and NEMHAUSER, G.L. (1978), "Optimality Conditions for a Dynamic Traffic Assignment Model", Transportation Science 12, 200-207.

ROBILLARD, P. (1974), "Multipath Assignment with Dynamic Input Flows". Transportation Research 8, 567-573

WARDROP, J.G. (1952), "Some Theoretical Aspects of Road Traffic Research", Proceedings of the Institut of Civil Engineers, Part II 1, 325-378.

YAGAR, S. (1970) "Dynamic Traffic Assignment by Individual Path Minimization and Queuing", Transportation Research 5, 179-196.

ZAWACK, D.J. and THOMPSON, G.L. (1987) "A Dynamic Space-Time Network Flow Model for City Traffic Congestion", Transportation Science 21, 153-162.